How can ACTS Work for you
Background Material - I

Tony Drummond
LBNL/NERSC

Motivation

http://acts.nersc.gov

Grand Challenges are ..fundamental problems in science and engineering, with potentially broad social, political, and scientific impact, that could be advanced by applying high performance computer resources

Office of Science and Technology

- Some grand challenges: electronic structure of materials, turbulence, genome sequencing and structural biology, global climate modeling, speech and language studies, pharmaceutical design, pollution, etc. .

Motivation

http://acts.nersc.gov

With the development of new kinds of equipment of greater capacity, and particularly of greater speed, it is almost certain that new methods will have to be developed in order to make the fullest use of this equipment. It is necessary not only to design machines for the mathematics, but also to develop a new mathematics for the machines - 1952, Hartree

- **Metropolis** grew out of physical chemistry in 1950's through attempts to calculate statistical properties of chemical reactions. Some areas of application: astrophysics, many areas engineering, and chemistry)
- **Fast Fourier Transform (FFT)**: implementation of Fourier Analysis. Some areas of application: image and signal processing, seismology, physics, radiology, acoustics and engineering)
- **Multigrids**: method for solving a wide variety of PDE. Some areas of application: physics, biophysics and engineering

Motivation

http://acts.nersc.gov

Computational science: can be characterized by the needs to gain understanding through the analysis of mathematical models using high performing performing computers

Community

- Scientists
- Engineers
- Mathematicians
- Economists, artists

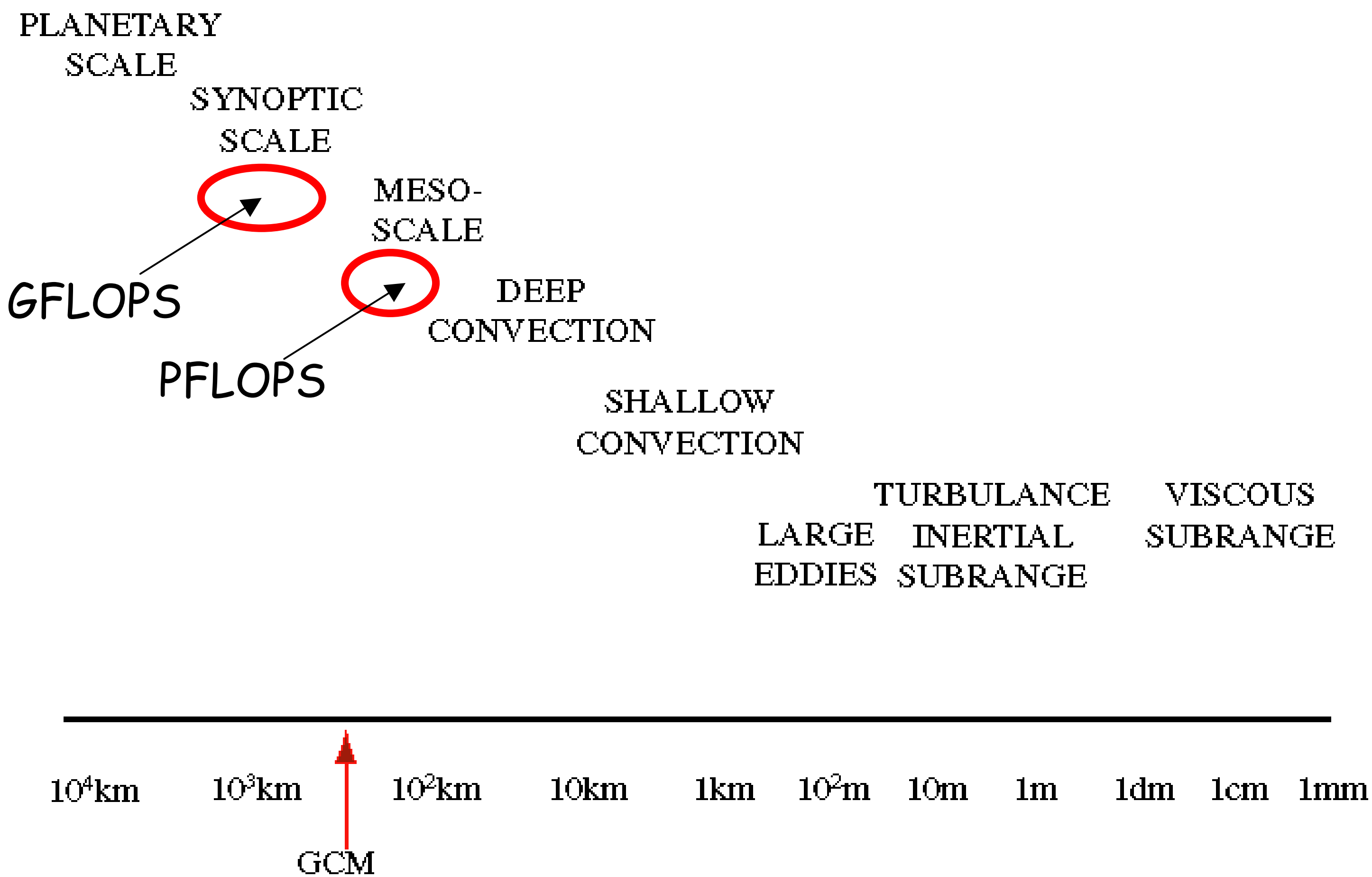
Multidisciplinary!

Computer Science	Mathematics:
Provides services ranging from Networking and visualization tools to algorithms	credibility of algorithms (error analysis, exact solutions, expansions, uniqueness proofs and theorems)

Motivation - Example I

http://acts.nersc.gov

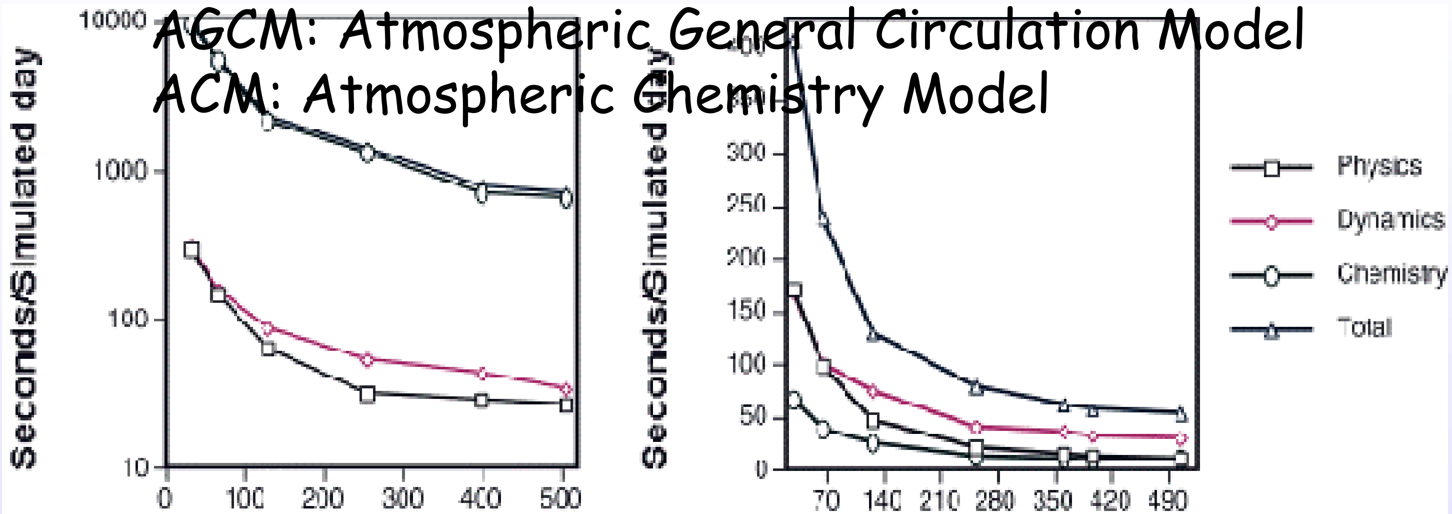
SPECTRUM OF ATMOSPHERIC PHENOMENA



Motivation - Example II

http://acts.nersc.gov

AGCM: Atmospheric General Circulation Model
ACM: Atmospheric Chemistry Model



AGCM/ACM
2.5 long x 2 lat, 30 layers
25-chemical species

AGCM/ACM
2.5 long x 2 lat, 30 layers
2-chemical species

- **Non-linear demand for resources (CPU - Memory)**
- **Multi-disciplinary application is more demanding**

The Hardware

Flynn's Taxonomy

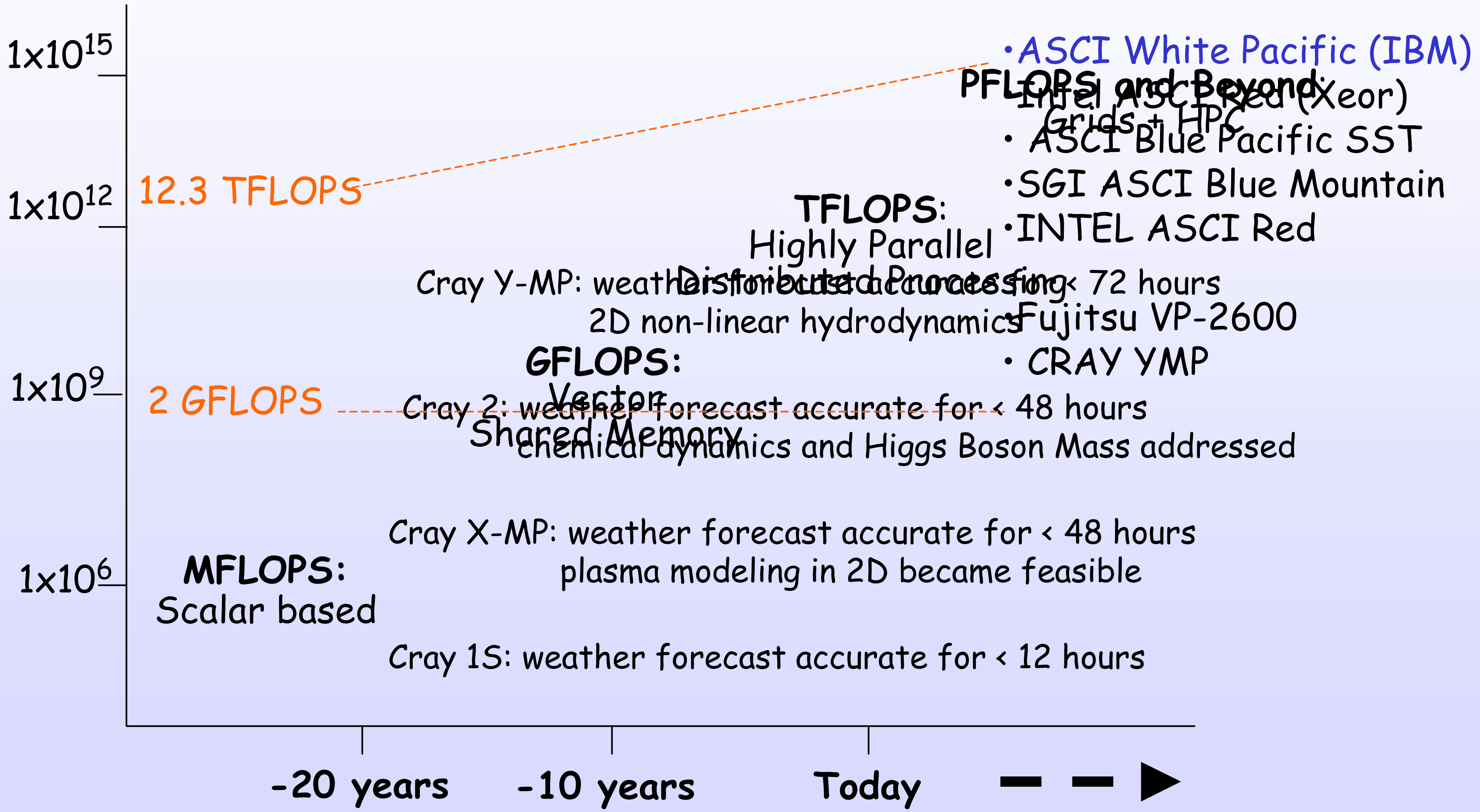
http://acts.nersc.gov

	Single Data Stream	Multiple Data Stream
Single Instruction Stream	**SISD**	**SIMD**
Multiple Instruction Stream	**MISD**	**MIMD**

High Performance Computers

http://acts.nersc.gov

FLoating Point Operations/Second (FLOPS)



The GRID

http://acts.nersc.gov

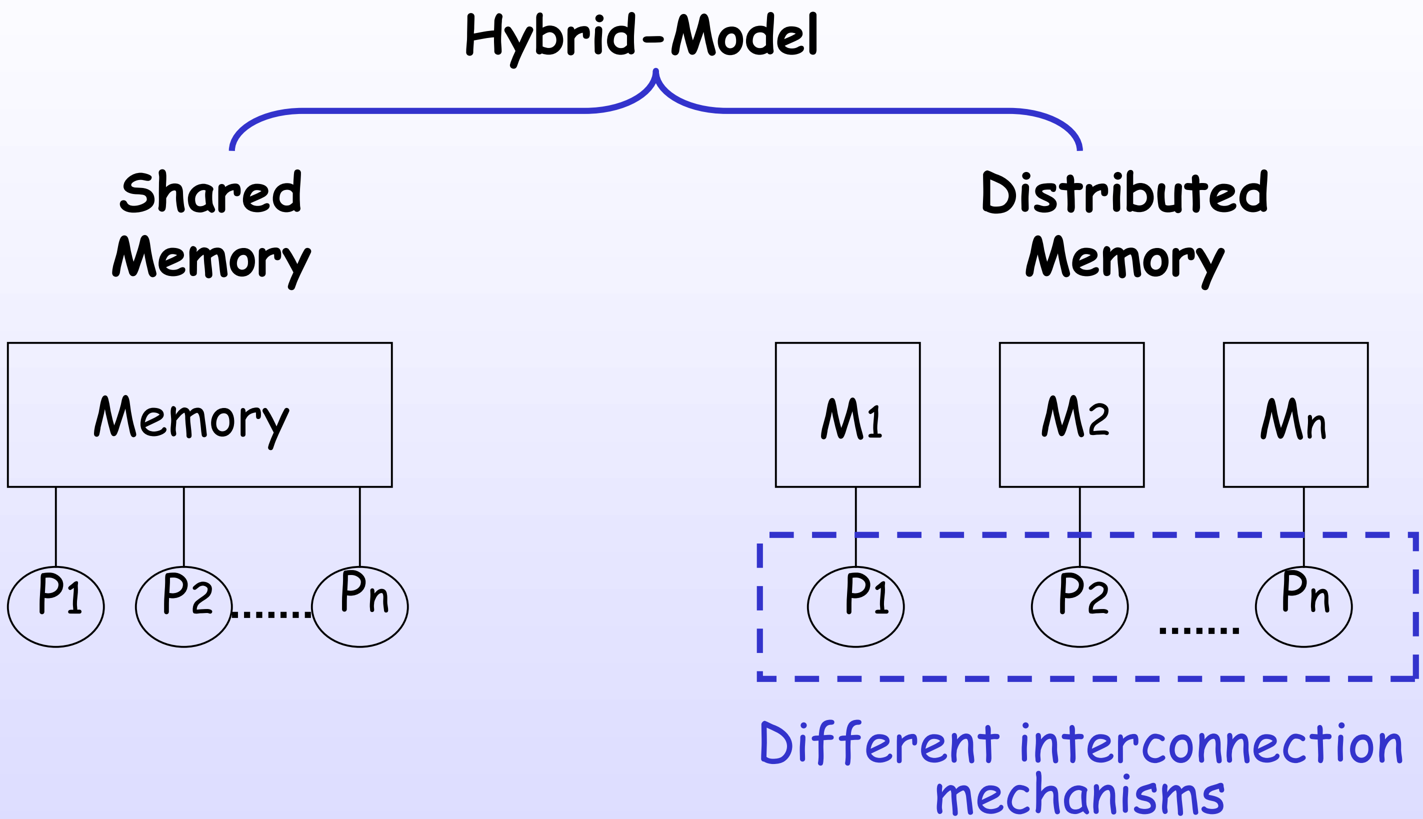
- A large pool of resources
  - Computers
  - Networks
  - Software
  - Databases
  - Instruments
  - people

Requirements from GRID implementation:

- Ubiquitous: ability to interface to the grid at any point and leverage whatever is available
- Resource Aware: manage heterogeneity of resources
- Adaptive: tailored to obtain maximum performance from resources

Shared vs. Distributed Memory

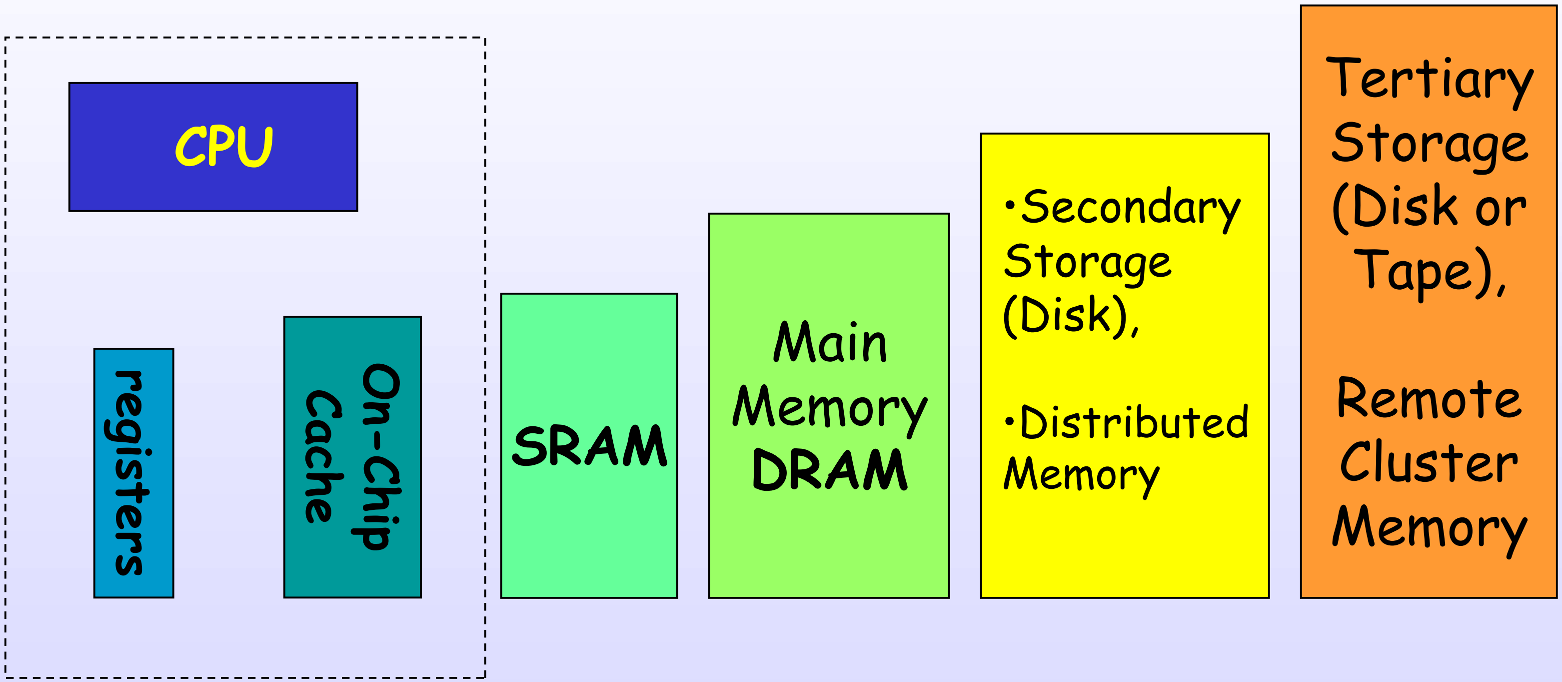
http://acts.nersc.gov



Memory Hierarchy

http://acts.nersc.gov

- *Where is the data? Why is data locality important?*



Speed	1's ns	10's ns	100's ns	.1's -10's ms	10's s
Size	100's bytes	Kbytes	Mbytes	Gbytes	Tbytes

Using the hardware

Levels of Parallelism

- Job and Task Level : Highest level of parallelism. Multidisciplinary applications running on a single computational resource or a collection of heterogeneous ones.

- Program Level: A single program and/or data is broken down into constituent parts

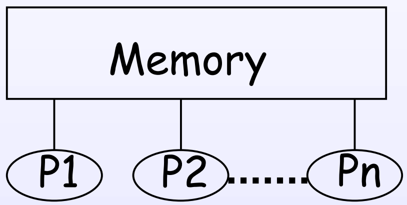
- Instruction Level: Pipeline and data streams

- Arithmetic and bit Level: Lowest level- CPU level

Parallel Programming Paradigms

http://acts.nersc.gov

Shared Memory



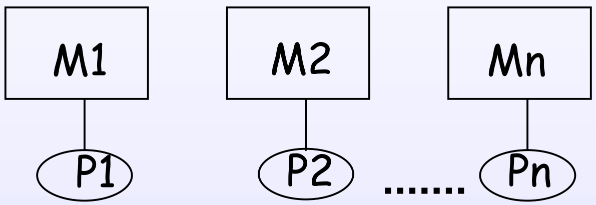
- easier to implement
- shared memory space
- mutual exclusion, contention

- shared area is use for sending and receiving data

- Data parallelism

- Message Passing

Distributed Memory



- virtual shared memory
- data is implicitly available to all

- Implicit mutual exclusion
- Only explicit synch
- Depends on Memory Hierarchy and Network

CPU vs. DRAM Performance

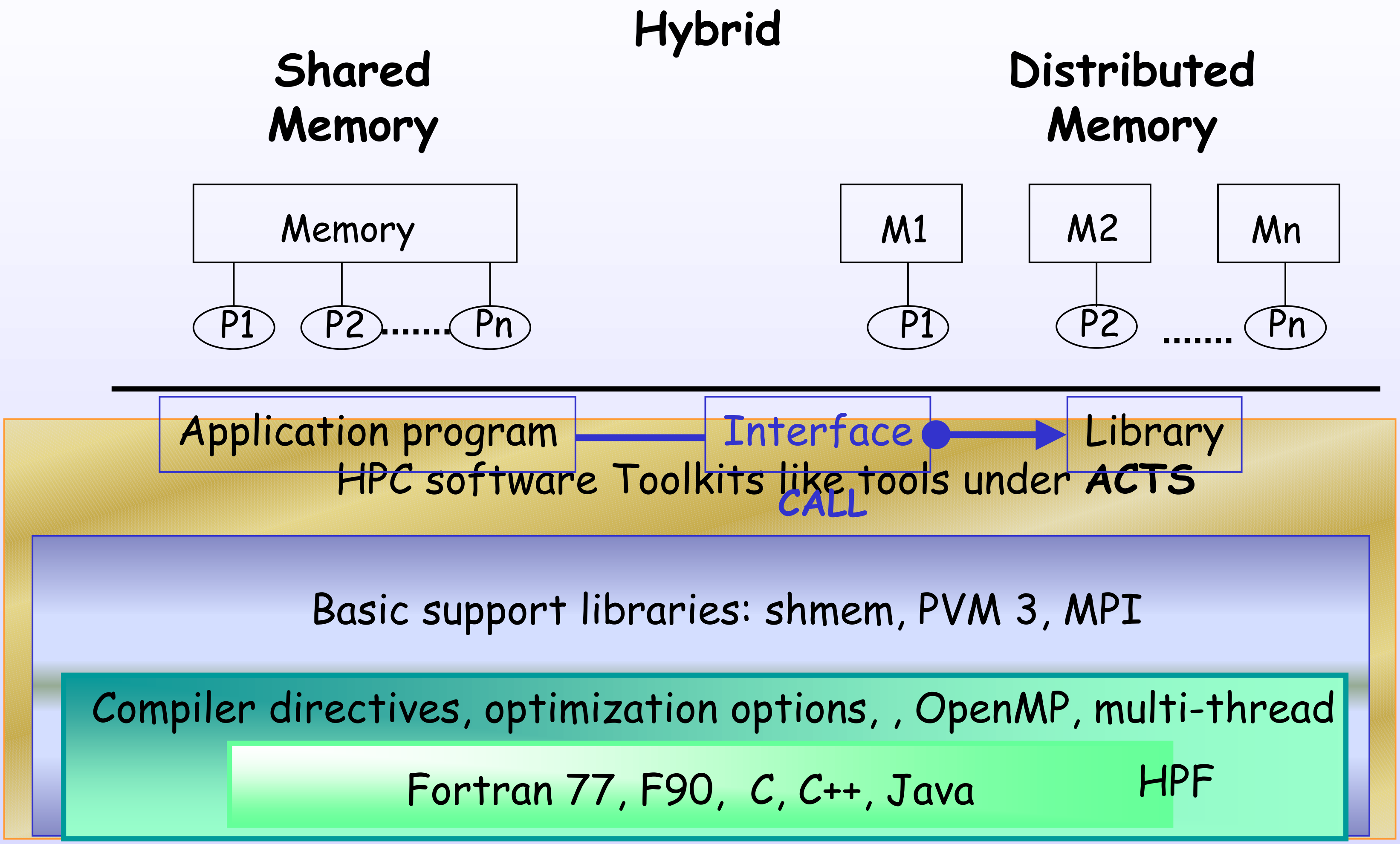
- Since 1980's, μProcs performance has increased at a rate of almost 60%/year

- Since 1980's, DRAM (latency) has improved at a rate of almost 9%/year

- Software required to bridge this gap
  - Tuned or optimized to existing hardware capabilities
  - Handle user needs (computational sciences)
  - Portable + Interoperable

Some Parallel Programming Tools

http://acts.nersc.gov



ACTS Tools Categorization

http://acts.nersc.gov

- Numerical
  - software that implements numerical algorithms
- Structural (“frameworks”)
  - software that manages data, communication
- Infra-structural
  - runtime, support tools, developer’s bag

Some Numerical Tools

http://acts.nersc.gov

- **Aztec**: iterative methods for solving sparse linear systems Tuminaro
- **Hypre**: collection of advanced preconditioners Falgout
- **PETSc**: methods for the solution of PDE related problems Gropp, and Balay
- **ScaLAPACK**: dense linear algebra computations Marques
- **SuperLU**: direct methods for sparse linear systems Li

- **TAO:** Toolkit for Advanced optimization More' and Benson

Some Structural (Frameworks)

http://acts.nersc.gov

- **Global Arrays**: portable, distributed array library, shared memory style of programming Nieplocha
- **Overture**: library of grid functions which derives from P++ arrays Quinlan

Infra-structural

http://acts.nersc.gov

- **CUMULVS** (Collaborative User Migration User Library for Visualization and Steering) Kohl
- **Globus**: infrastructure for high performance distributed computing. Bag of services for the grid Czajkowski
- **PADRE** (Parallel Asynchronous Data and Routing Engine) : abstracts the details of representing and managing distributed data Quinlan
- **PAWS** (Parallel Application WorkSpace) provides interapplication support in heterogeneous computing environments Rasmussen

Infra-structural (cont.)

http://acts.nersc.gov

- **SILOON** (Scripting Interface Languages for Object-Oriented Numerics): scripting features   Rasmussen
- **TAU** (Tuning and Analysis Utilities): advanced performance analysis and tuning   Malony

acts-support@nersc.gov

http://acts.nersc.gov